SciDataCon 2016

Conference paper

A Portfolio Approach to a Sustainable Business Model for Scientific Data Stewardship

Abstract

Examination of sustainable business models is conducted to identify opportunities for ensuring continuing support of scientific data centers and other digital repositories that are responsible for the stewardship of resources to support science and scholarship. Based on a typology of sustainable approaches for scientific data stewardship, a portfolio approach is proposed for repositories that are seeking to improve the sustainability of their business models. Such a diversified approach should not only support current operational services but also reduce the risks to long-term data stewardship.

Introduction

Continuing stewardship is needed for scientific data that have been determined to be valuable for future use. Considering the ongoing costs of providing continuing stewardship for scientific data, a sustainable approach is needed to ensure ongoing operation of the digital repositories, such as scientific data centers, that are responsible for providing stewardship for scientific data of enduring value. Increasing operational efficiency and selective acquisition of data holdings can help to manage costs and resource utilization over time, but resources also are necessary to provide the infrastructure and support ongoing efforts needed to maintain the accessibility, integrity, and usability of scientific data in the long run. Examination of sustainable business models offers an opportunity to identify resources for enabling continuing support of scientific data centers and other digital repositories that are responsible for the stewardship of scientific data.

Sustainable Business Models for Scientific Data Stewardship

Business models based on discrete revenue streams or cooperative models can contribute to sustainable scientific data stewardship (Downs and Chen, 2012). Differentiating between discrete revenue stream models and cooperative models enables scientific data centers and other digital repositories to diversify among these types of opportunities when choosing business models to support their operations. As depicted in **Figure 1**, a typology of sustainable approaches for scientific data stewardship offers insight into the variety of business models that can be adopted for continuous support for scientific data (Downs and Chen, 2013). Diversifying among the types of business models available for sustainable support of a scientific data center should reduce risks of gaps in operational and infrastructure support by reducing the reliance on a single business model and enabling the pursuit of alternative options for support (Chen and Downs, 2014). The selection of opportunities from the combination of available business models to be pursued by a

particular scientific data center should be informed by an assessment of current and long-term needs as well as by an assessment of the opportunities that are available to the data center.

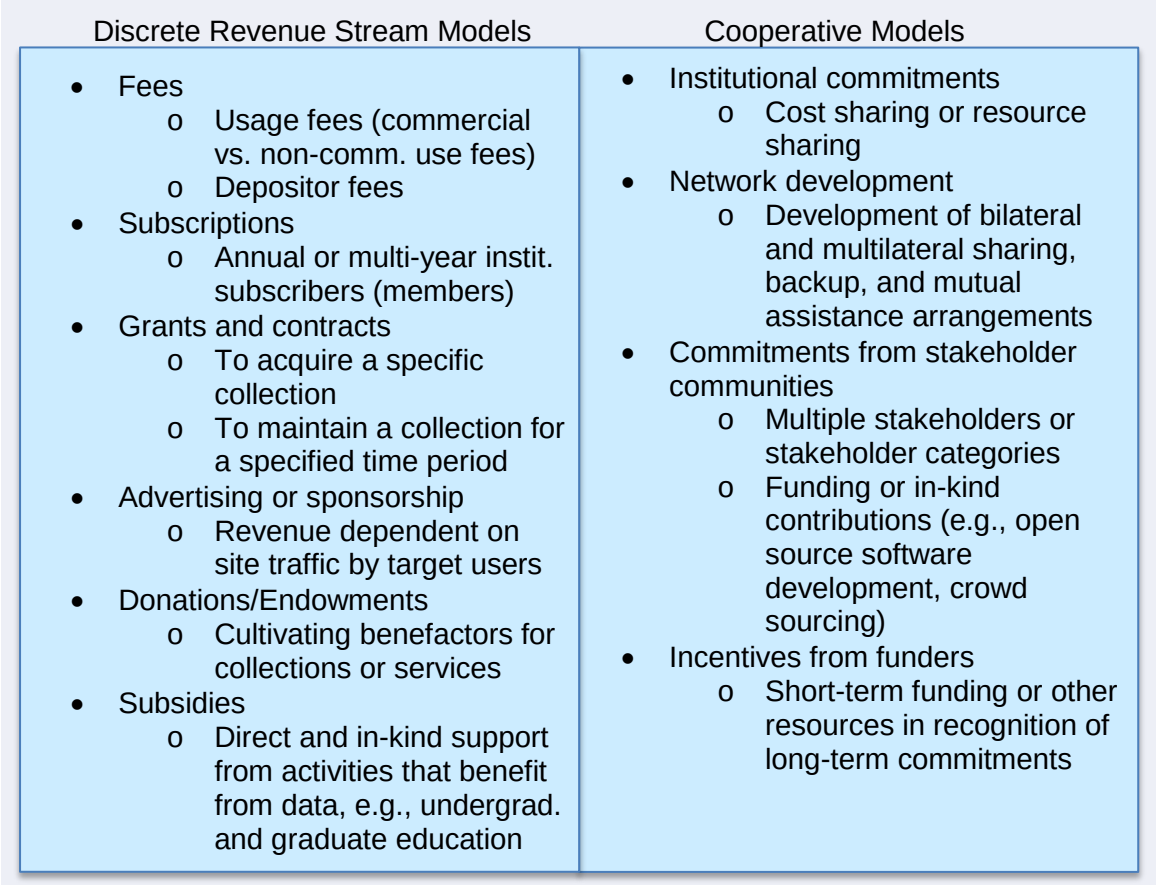

Figure 1: Typology of Sustainability Approaches for Scientific Data Stewardship. Adapted from Downs and Chen (2013).

In addition to diversifying among the types of approaches selected to achieve a sustainable business model for operation of a scientific data center, it is important to consider longitudinal aspects of adopting such approaches. The criteria for selecting current approaches for support of a scientific data center may be different than the criteria for selecting long-term approaches for such support. Therefore, a portfolio approach is recommended for selecting a business model for a repository that provides data stewardship and access services to enable science and scholarship. In applying a portfolio approach, an assessment is conducted when considering the options for ensuring current support and a separate assessment is conducted when considering the options for ensuring future support. Identifying current opportunities for support is necessary and critical for maintaining current operations and for developing new business opportunities. However, since the future availability of current opportunities cannot be predicted with absolute certainty, alternative future opportunities also should be considered. If the current business model is considered stable for the foreseeable future, then alternative opportunities can be identified as part of contingency plans to ensure that future options are available. The adoption of a business model portfolio that utilizes contingency plans has implications for the goals, organizational structure, levels of services, and sources of support of the data center and how it prioritizes resources for core data stewardship activities versus other data-related activities.

A Sustainable Business Model Portfolio for a Scientific Data Center

Identifying opportunities for ensuring both current support and long-term support can contribute to a business model portfolio for supporting the continuing stewardship of scientific data. The National Aeronautics and Space Administration (NASA) Socioeconomic Data and Applications Center (SEDAC) serves as a case study to begin exploring such a portfolio approach to the adoption of sustainable business models for scientific data stewardship. In its business model to support its current active archive operations, SEDAC primarily receives support from NASA, with some shared support across other sponsored projects operated by the Center for International Earth Science Information Network (CIESIN) at Columbia University. Structural support in the form of such grants and contracts offers advantages, including operating as part of the distributed architecture of the NASA Earth Observing System Data and Information System (EOSDIS). Operating within the EOSDIS also facilitates support from within the system and from among its centers to foster continuing improvements for data stewardship, management, and use (Ramapriyan et al., 2010). Recognizing that long-term stewardship of scientific data also is needed and that support from grants and contracts is not guaranteed forever, SEDAC also has established contingency plans, primarily from institutional commitments but potentially also from stakeholder commitments, to ensure continuing support for its Long-Term Archive if a lapse in funding should occur (Chen and Downs, 2014; Downs and Chen, 2010; 2012). The initial exploration of a portfolio approach to a sustainable business model for scientific data stewardship of NASA SEDAC data is illustrated in **Figure 2**, with primary sources of support identified. Additional alternative options for ensuring continuing support of the Long-Term Archive also need to be explored as well as possible tradeoffs, such as the initiation of fees versus open access. Comparisons of approaches, such as membership models that have been employed successfully, also could inform exploration into the feasibility of adopting such approaches for long-term data stewardship.

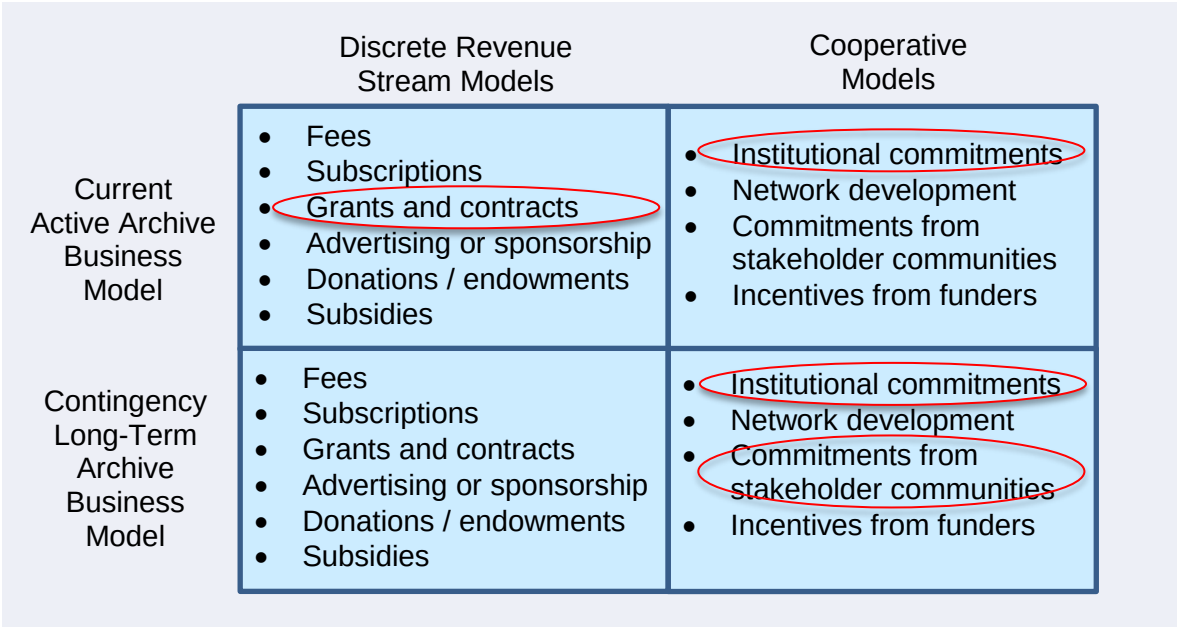

Figure 2: A Portfolio Approach to a Sustainable Business Model for Scientific Data Stewardship.

Conclusions

Based on a typology of sustainable approaches for scientific data stewardship, a portfolio approach to adopting sustainable business models is proposed for scientific data centers and other repositories to ensure long-term stewardship of their holdings. Adoption of the portfolio approach for a scientific data center is initially explored to examine how such an approach can reduce potential risks to long-term data stewardship. Developing a portfolio approach provides scientific data centers and other digital repositories with an opportunity to plan beyond the current state of sustainability to consider alternative opportunities for possible future states. Contingency planning and the exploration of possible alternative models offers a way to identify available alternative opportunities that may or may not be needed to ensure future sustainability for data repositories in the event of circumstances that cannot be predicted. If alternative opportunities are needed, contingency plans and feasible alternatives can be adopted as part of a business model portfolio to reduce risks and ensure continuity of the services that are necessary to provide sustainable stewardship of scientific data.

Acknowledgements

The authors appreciate the support received from the National Aeronautics and Space Administration (NASA) under Contract NNG13HQ04C for the Socioeconomic Data and Applications Distributed Active Archive Center (DAAC).

Competing Interests

The authors are employed by the Center for International Earth Science Information Network (CIESIN), of Columbia University, which operates the NASA Socioeconomic Data and Applications Center (SEDAC).

References

Chen RS and **Downs RR** 2014 Government-Academic Partnership in Hosting an Interdisciplinary Data Center. SciDataCon 2014, New Delhi, India, 2-5 November 2014.

Downs RR and **Chen RS** 2010 Self-Assessment of a Long-Term Archive for Interdisciplinary Scientific Data as a Trustworthy Digital Repository. *Journal of Digital Information*, 11(1): Open Repositories 2009 Special Issue. DOI: http://journals.tdl.org/jodi/article/view/753

Downs, R R and **Chen, R S** 2012 Towards Sustainable Stewardship of Digital Collections of Scientific Data. In: GSDI World Conference (GSDI 13) Conference Proceedings, Québec City, Canada, 14-17 May 2012. Available online from Columbia University Academic Commons. DOI: http://dx.doi.org/10.7916/D8KK9BTG

Downs, R R and **Chen, R S** 2013 Comparing Approaches for the Sustainability of Scientific Data Repositories, Columbia University Academic Commons. DOI: http://hdl.handle.net/10022/AC:P:19169

Ramapriyan, H K, Behnke, J, Sofinowski, E, Lowe, D and **Esfandiari, M A** 2010 Evolution of the earth observing system (EOS) data and information system (EOSDIS). In Di, L and Ramapriyan, H K *Standard-Based Data and Information Systems for Earth Observation*. Berlin Heidelberg: Springer. pp. 63-92. DOI: http://dx.doi.org/10.1007/978-3-540-88264-0_5